

RESEARCH ARTICLE

Open Access

# Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface

Alessandro Siglioccolo<sup>†</sup>, Alessandro Paiardini<sup>†</sup>, Maria Piscitelli and Stefano Pascarella<sup>\*</sup>

## Abstract

**Background:** Halophiles are extremophilic microorganisms growing optimally at high salt concentrations. There are two strategies used by halophiles to maintain proper osmotic pressure in their cytoplasm: accumulation of molar concentrations of potassium and chloride with extensive adaptation of the intracellular macromolecules ("salt-in" strategy) or biosynthesis and/or accumulation of organic osmolytes ("osmolyte" strategy). Our work was aimed at contributing to the understanding of the shared molecular mechanisms of protein haloadaptation through a detailed and systematic comparison of a sample of several three-dimensional structures of halophilic and non-halophilic proteins. Structural differences observed between the "salt-in" and the mesophilic homologous proteins were contrasted to those observed between the "osmolyte" and mesophilic pairs.

**Results:** The results suggest that haloadaptation strategy in the presence of molar salt concentration, but not of osmolytes, necessitates a weakening of the hydrophobic interactions, in particular at the level of conserved hydrophobic contacts. Weakening of these interactions counterbalances their strengthening by the presence of salts in solution and may help the structure preventing aggregation and/or loss of function in hypersaline environments.

**Conclusions:** Considering the significant increase of biotechnology applications of halophiles, the understanding of halophilicity can provide the theoretical basis for the engineering of proteins of great interest because stable at concentrations of salts that cause the denaturation or aggregation of the majority of macromolecules.

## Background

Organisms thriving in "extreme environments", such as thermophiles, alkalophiles, acidophiles, halophiles, piezophiles and psychrophiles, have drawn much interest in the scientific community because of the molecular adaptation they underwent during evolution and for their biotechnological potential [1-4]. The environmental challenges that extremophilic organisms have to face necessitate, besides other physiological modifications, biosynthesis of macromolecules stable and active at environmental physical-chemical extreme conditions. These macromolecules display clearly distinguished features when compared to the macromolecules from

microorganisms found in the "normal" (mesophilic) environments [5].

Halophilic microorganisms are salt-loving extremophilic organisms that grow optimally at high salt concentrations. They were found [6] mainly in marine salterns and hypersaline lakes, such as the Great Salt Lake and the Dead Sea. A survey of the salt requirements in the microbial world shows a continuum of properties which makes it very difficult to define by sharp limits what a halophilic microorganism is. The accepted view [6] distinguishes the halophilic organisms in: extreme halophiles (growing best in media containing 2.5-5.2 M salt), borderline extreme halophiles (growing best in media containing 1.5-4.0 M salt), moderate halophiles (growing best in media containing 0.5-2.5 M salt), and halotolerant microorganisms that do not show an absolute requirement for salt for growth but grow well up to often very high salt concentrations (considered

\* Correspondence: Stefano.Pascarella@uniroma1.it

† Contributed equally

Dipartimento di Scienze Biochimiche "A. Rossi Fanelli", Università di Roma La Sapienza, 00185 Roma, Italy

extremely halotolerant if the growth range extends above 2.5 M salt).

Distribution of halophilic microorganisms within the tree of life shows that they are widespread in the bacterial and *Archaea* kingdoms. Eukaryotic halophilic microorganisms, such as fungi and algae, are also known [7].

All halophilic microorganisms share a basic property: their cytoplasm must be at least isoosmotic with their surrounding medium. There are two strategies used by halophilic microorganisms to maintain proper osmotic pressure in their cytoplasm. The first involves accumulation of molar concentrations of potassium and chloride. This strategy requires extensive adaptation of the intracellular enzymatic machinery to the presence of salts, as the proteins should maintain their proper conformation and activity at near-saturating salt concentrations. The proteome of such organisms is highly acidic, and most proteins denature when suspended in low salt. Generally, the microorganisms relying on such "salt-in" strategy are obligate halophilic *Archaea*. The other strategy of haloadaptation is based on the biosynthesis and/or accumulation of organic osmotic solutes (osmolytes) such as ectoin, glycine betaine or others [8]. Cells relying on this strategy exclude salts from their cytoplasm as much as possible. The high concentrations of organic "compatible" solutes do not greatly interfere with normal enzymatic activity. Fewer adaptations of the cells' proteome are therefore needed. Such organisms can often adapt to a broad salt concentration range [6].

Halophilic proteins from "salt-in" organisms are distinguished from their non-halophilic homologous proteins by exhibiting remarkable instability in low salt concentration and by maintaining soluble and active conformations in an environment generally detrimental to other proteins. Indeed, hypersaline conditions favor protein aggregation and collapse, interfere with the electrostatic interactions between protein residues, and are responsible for a general decrease in the availability of water molecules [9]. Halophilic proteins, rather than being unfolded by these conditions, appear to be dependent on the presence of salts. In recent years, detailed investigations have tried to unveil the relationships between structure and stability in halophilic proteins [10]. In particular, these studies suggested that the halophilic proteins bind significant amounts of salt and water in solvent conditions similar to their physiological environment. The peculiar ability of halophilic proteins to bind large amount of salts is largely dependent on the number of acidic amino acids on protein surface [11-17]. The role of electrostatic interactions in the stability and folding of halophilic proteins has been investigated and recognized as a key determinant of haloadaptation [18,19]. Moreover, it was observed that halophilic proteins are characterized by a general

decrease in hydrophobic amino acid frequency and a greater propensity to form random-coil structures, rather than  $\alpha$ -helices [15]. Indeed, protein folding and adequate stability of the native structure in a hypersaline environment may require evolutionary modulation of the hydrophobic interactions occurring at the protein core. Most of the studies carried out to unveil the structural characteristics of halophilic proteins, were based on sequence comparison at proteome and genome levels or were focused onto single or few protein families. These analyses provided undoubtedly valuable indications on the biophysical and biochemical properties of the halophilic proteins. However, scrutiny of proteome and genomic sequences may not unravel subtle differences at the three-dimensional structural level while structural analysis of a single or few protein families may lack sufficient generalization. For these reasons we report in this work a systematic comparison between the available three-dimensional structures of halophilic enzymes deposited in the data banks and the structure of one of their homologues, to investigate the differences possibly related to shared strategies of structural adaptation to high salt environments. Use of three-dimensional structure made it possible to investigate subtle modifications of the surface and hydrophobic core of the halophilic proteins especially at the level of conserved hydrophobic contacts [20].

## Results

### Data set

A set of 15 halophilic enzymes, 9 of which from "extreme" halophiles and 6 from "halotolerant" organisms, was collected along with their non-halophilic structural homologues (Table 1). Among the halophilic enzymes, 8 come from *Archaea* that adopt the "salt-in" strategy (SALTIN), and 7 from *Eubacteria* that adopt the "osmolytes" strategy (OSMOL). In this paper, the entire analysis was carried out considering the two groups separately.

Four pairs of halophilic enzymes (PDB ID: 2B5W, 2J5K, 1NML, 3IGN) have a quaternary structure different from their non-halophilic homologs (Table 1) and therefore were excluded from the calculations involving proteins in their quaternary structure, such as calculation of the electrostatic potential surface. The pair cytochrome  $c_{552}$  from *Marinobacter hydrocarbonoclasticus* (PDB ID: 1CNO) and from *Pseudomonas stutzeri* (PDB ID: 1ETP) has been included in the analysis, for the reasons explained in Methods section.

### Accessible surface area and electrostatic potential

To calculate the relative solvent accessibility (ASA) and the magnitude of the relative polar and apolar components, protein structures were used in their quaternary

**Table 1 Data set.**

	PDB <sup>a)</sup>	Q <sup>b)</sup>	Organism <sup>c)</sup>	[NaCl] <sup>d)</sup>	Survival <sup>e)</sup>	Res (Å) <sup>f)</sup>	Name <sup>g)</sup>	PDB <sup>a)</sup>	Q <sup>b)</sup>	Organism <sup>c)</sup>	Res (Å) <sup>f)</sup>	Sequence identity(%) <sup>h)</sup>
1	1DOI	1	<i>Haloarcula marismortui</i> (A)	3.4-3.9M	Salt-in	1.90	Ferredoxin	1FXA	1	<i>Anabaena sp.</i> (B)	2.50	51%
2	1NWZ	1	<i>Halorhodospira halophila</i> (B)	1.5-5.1M	Osmolytes	0.82	Photoreceptor	1MZU	1	<i>Rhodospirillum centenum</i> (B)	2.00	46%
3	1TJO	12	<i>Halobacterium salinarum</i> (A)	3.9M	Salt-in	1.60	DNA-protecting protein	2VXX	12	<i>Thermosinechococcus elongatus</i> (B)	2.40	36%
4	2B5W	2	<i>Haloferax mediterranei</i> (A)	3.4-4.3M	Salt-in	1.60	Glucose dehydrogenase	2CD9	4	<i>Sulfolobus solfataricus</i> (A)	1.80	30%
5	2CC6	12	<i>Halobacterium salinarum</i> (A)	3.9M	Salt-in	1.27	Dodecin	2V18	12	<i>Thermus thermophilus</i> (B)	2.59	42%
6	3IBM	2	<i>Halorhodospira halophila</i> (B)	1.5-5.1M	Osmolytes	2.00	Cupin 2 domain-containing protein	3KGZ	2	<i>Rhodopseudomonas palustris</i> (B)	1.85	44%
7	1ITK	2	<i>Haloarcula marismortui</i> (A)	3.4-3.9M	Salt-in	2.00	Catalase-peroxidase	2FXG	2	<i>Burkholderia pseudomallei</i> (B)	2.00	60%
8	2AZ3	6	<i>Halobacterium salinarum</i> (A)	3.9M	Salt-in	2.20	Nucleoside diphosphate kinase	3B54	6	<i>Saccharomyces cerevisiae</i> (B)	3.10	54%
9	2J5K	4	<i>Haloarcula marismortui</i> (A)	3.4-3.9M	Salt-in	1.95	Malate dehydrogenase	1Y6J	2	<i>Clostridium thermocellum</i> (B)	3.01	33%
10	1CNO	2	<i>Marinobacter hydrocarbonoclasticus</i> (B)	0.6-0.85M	Osmolytes	2.20	Cytochrome c <sub>552</sub>	1ETP	1	<i>Pseudomonas stutzeri</i> (B)	2.20	47%
11	1NML	1	<i>Marinobacter hydrocarbonoclasticus</i> (B)	0.6-0.85M	Osmolytes	2.20	Cytochrome c peroxidase	3HQ6	2	<i>Geobacter sulfurreducens</i> (B)	2.20	64%
12	2VPN	1	<i>Halomonas elongata</i> (B)	0.5-1.4 max 5.5M	Osmolytes	1.55	Prisplasmic ectoin-binding protein	3FXB	1	<i>Ruegeria pomeroyi</i> (B)	2.90	62%
13	3IFV	3	<i>Haloferax volcanii</i> (A)	1.7M	Salt-in	2.00	Proliferating cell nuclear antigen	1RWZ	3	<i>Archaeoglobus fulgidus</i> (A)	1.80	36%
14	3IGN	1	<i>Marinobacter aquaeleioi vt8</i> (B)	0.6-0.85M	Osmolytes	1.83	GGDEF domain	3ISC	2	<i>Pseudomonas aeruginosa pao1</i> (B)	1.94	40%
15	3BSM	8	<i>Chromohalobacter salexigens</i> (B)	1.2-1.7M	Osmolytes	2.20	D-mannonate dehydratase	2QJJ	8	<i>Novosphingobium aromaticivorans</i> (B)	1.80	66%

List of protein pairs utilized in the work. Boldfaced PDB codes indicate extreme halophiles.

<sup>a)</sup> Protein Data Bank code

<sup>b)</sup> Number of monomers in the biological unit

<sup>c)</sup> Source organisms. (A) = Archaea; (B) = Bacteria

<sup>d)</sup> Optimal range of NaCl concentration for growth. The first 9 lines contain structures from extreme halophiles; the remaining 6 lines report moderate halophiles.

<sup>e)</sup> The survival strategy adopted by corresponding organism: "salt-in" or "osmolytes" strategy

<sup>f)</sup> Crystallographic resolution

<sup>g)</sup> Protein name as reported in the Protein Data Bank

<sup>h)</sup> Sequence percentage of identity between protein pairs

form. The surface polar area formed by oxygen and nitrogen atoms was calculated. The area subset formed only by side-chain oxygen or nitrogen atoms (namely, excluding atoms belonging to the polypeptide backbone) was also considered.

In SALTIN enzymes, the fractional apolar component of the relative solvent accessible surface is significantly smaller than the corresponding area in non-halophilic counterparts ( $\Delta\text{ApA}$  in Table 2). The polar area formed by oxygen atoms is significantly larger in halophilic enzymes; the same result is obtained when the analysis is repeated excluding backbone oxygen atoms and including only side-chain oxygen atoms. Likewise, the

polar area formed by nitrogen atoms is significantly smaller than the corresponding area in non-halophilic counterparts; the same result is obtained when the analysis is repeated without backbone nitrogen atoms and including only side-chain nitrogen atoms. These differences are reflected in the surface electrostatic potential, which is significantly more negative in halophilic enzymes than in the non-halophilic counterparts (Table 3). In OSMOL sample, the apolar component of the relative solvent accessibility was also significantly smaller than the corresponding area in non-halophilic counterparts (Table 2); however, the decreased surface is caused solely by the increase of the polar area formed by side-

**Table 2 ΔASA in the SALTIN and OSMOL samples.**

SALTIN HALOPHILES	NON HALOPHILES	ΔApA <sup>a)</sup>	ΔTot O <sup>b)</sup>	ΔSc O <sup>c)</sup>	ΔTot N <sup>d)</sup>	ΔSc N <sup>e)</sup>	OSMOL HALOPHILES	NON HALOPHILES	ΔApA <sup>a)</sup>	ΔTot O <sup>b)</sup>	ΔSc O <sup>c)</sup>	ΔTot N <sup>d)</sup>	ΔSc N <sup>e)</sup>
1DOI	1FXA	-0.07	0.06	-0.08	0.01	0.01	1NWZ	1MZU	-0.05	0.08	0.05	-0.03	-0.02
1TJO	2VXX	-0.09	0.14	0.15	-0.04	-0.04	3IBM	3KGZ	-0.07	0.02	0.02	0.04	0.04
2CC6	2V18	-0.07	0.17	0.15	-0.10	-0.09	1CNO	1ETP	-0.02	0.00	0.01	0.01	0.02
1ITK	2FXG	-0.10	0.13	0.14	-0.03	-0.03	2VPN	3FXB	-0.06	0.06	0.05	-0.01	-0.00
2AZ3	3B54	-0.02	0.09	0.08	-0.06	-0.06	3BSM	2QJJ	0.00	0.00	0.00	0.01	0.01
3IFV	1RWZ	-0.03	0.07	0.08	-0.04	-0.03							
Total <sup>f)</sup>		-0.38	0.66	0.52	-0.28	-0.25			-0.21	0.16	0.14	0.02	0.05
Average <sup>g)</sup>		-0.06	0.11	0.90	-0.05	-0.04			-0.04	0.03	0.03	0.00	0.01
t-test <sup>h)</sup>		<b>0.00</b>	<b>0.00</b>	<b>0.05</b>	<b>0.02</b>	<b>0.03</b>			<b>0.03</b>	0.11	<u>0.07</u>	0.71	0.42
Wilcoxon <sup>h)</sup>		<b>0.03</b>	<b>0.03</b>	<b>0.05</b>	<b>0.05</b>	<b>0.05</b>			<b>0.04</b>	0.14	<b>0.04</b>	0.50	0.50

Differences of fractional accessibility surface area (ΔASA) in the SALTIN and OSMOL samples for different class of atoms. The differences are between the surface areas calculated in the halophilic protein and the corresponding areas in the non-halophilic counterpart. The calculations were performed considering the proteins in their quaternary structure.

- a) Apolar ΔASA difference between fractional apolar exposed areas of the halophilic protein and the corresponding non-halophilic homolog
- b) Oxygen atom fractional ΔASA
- c) Side-chain oxygen atom fractional ΔASA
- d) Nitrogen atom fractional ΔASA
- e) Side-chain nitrogen fractional ΔASA
- f) Total fractional ΔASA
- g) Average fractional ΔASA
- h) Boldfaced and underlined *p*-values indicate significant or possible trend, respectively

chain oxygen atoms. In fact, there is no significant difference in the areas formed by oxygen atoms or nitrogen atoms. The analysis was repeated considering only the subset of residues identically conserved between the halophilic and the corresponding non-halophilic homologue in the SALTIN and OSMOL samples. In this case, no significant difference was seen in the exposed areas (see additional file 1).

Surface electrostatic potential displays no significant difference although, in 3 out of 5 pairs, it is more negative in the OSMOL halophilic proteins (Table 3).

#### Apolar contact area

The apolar contact area (ACA) between residues was calculated for three different structural regions namely the core, the interface between the monomers and the protein surface. Residues were assigned to the various structural regions according to the alternative accessibility thresholds described in the Methods section.

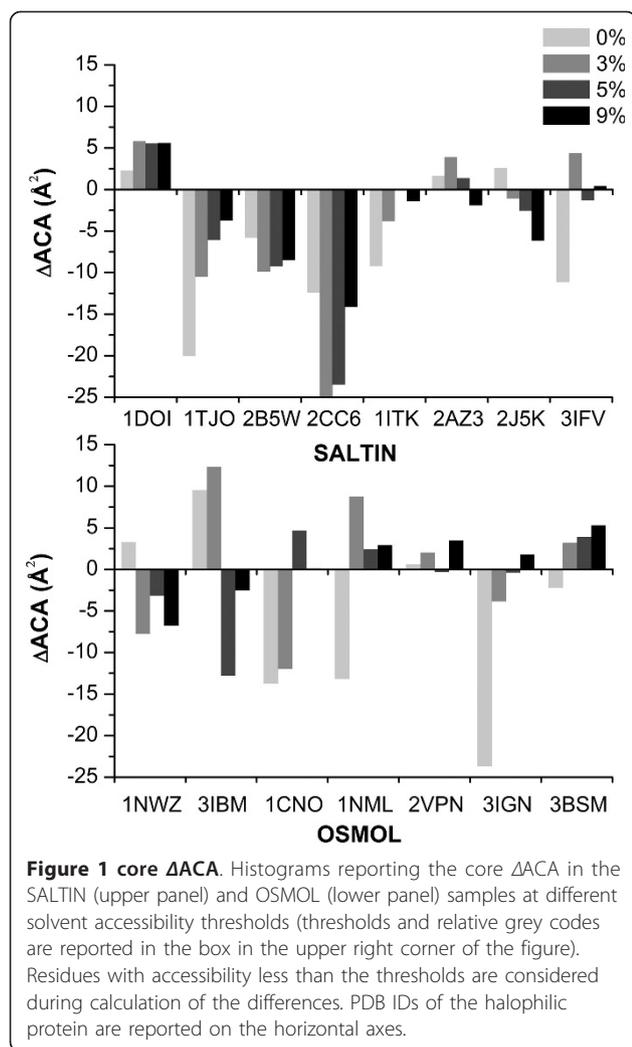
ACA in the core of SALTIN halophilic proteins is not significantly different, according to the statistical tests, from the ACA of their homologous counterparts. However, it is consistently smaller (Figure 1). This trend is

**Table 3 ΔAAP for SALTIN and OSMOL samples.**

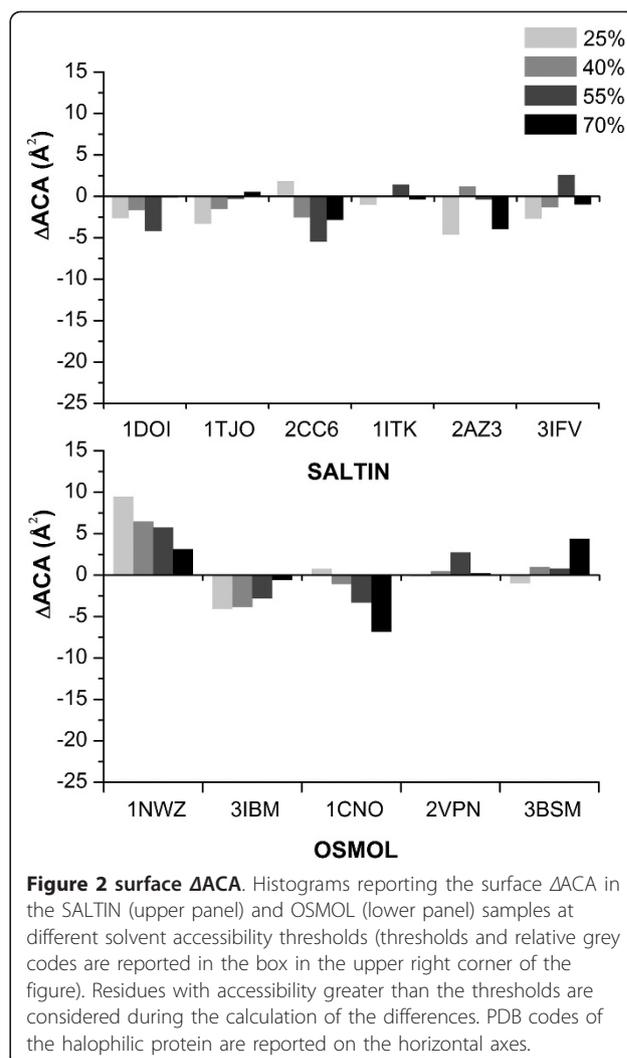
SALTIN HALOPHILES	NON HALOPHILES	ΔAAP	OSMOL HALOPHILES	NON HALOPHILES	ΔAAP
1DOI	1FXA	-5.96	1NWZ	1MZU	-2.91
1TJO	2VXX	-76.92	3IBM	3KGZ	-6.85
2CC6	2V18	-49.24	1CNO	1ETP	4.83
1ITK	2FXG	-25.62	2VPN	3FXB	-2.89
2AZ3	3B54	-32.92	3BSM	2QJJ	0.06
3IFV	1RWZ	-17.59			
Total <sup>a)</sup>		-208.26			-7.75
Average <sup>b)</sup>		-34.71			-1.55
t-test <sup>c)</sup>		<b>0.02</b>			0.47
Wilcoxon <sup>c)</sup>		<b>0.03</b>			0.50

Differences of average atomic potential (ΔAAP) for SALTIN and OSMOL samples. Units are kT/e.

- a) Total ΔAAP
- b) Average ΔAAP
- c) Boldfaced digits indicate significant *p*-values



**Figure 1 core  $\Delta ACA$ .** Histograms reporting the core  $\Delta ACA$  in the SALTIN (upper panel) and OSMOL (lower panel) samples at different solvent accessibility thresholds (thresholds and relative grey codes are reported in the box in the upper right corner of the figure). Residues with accessibility less than the thresholds are considered during calculation of the differences. PDB IDs of the halophilic protein are reported on the horizontal axes.



**Figure 2 surface  $\Delta ACA$ .** Histograms reporting the surface  $\Delta ACA$  in the SALTIN (upper panel) and OSMOL (lower panel) samples at different solvent accessibility thresholds (thresholds and relative grey codes are reported in the box in the upper right corner of the figure). Residues with accessibility greater than the thresholds are considered during the calculation of the differences. PDB codes of the halophilic protein are reported on the horizontal axes.

more pronounced when the core contains only residues at 0% accessibility. At variance, OSMOL enzymes have an ACA that is globally comparable with that of non-halophilic counterparts (Figure 1), even at 0% accessibility threshold.

Likewise, ACA in the SALTIN monomer interface is not significantly dissimilar from the ACA of their non-halophilic counterparts. However, even in this case, all the  $\Delta ACAs$  were negative except for the pair 3IFV-1RWZ. The analysis at interface for OSMOL enzymes is reduced to two pairs and is therefore of no statistical interest, however, the ACA formed by interfacial residues in halophiles is comparable to that of non-halophilic organisms.

The surface ACA in SALTIN and OSMOL halophilic enzymes is not significantly different from the corresponding surface in the non-halophilic counterparts (Figure 2).

Differences in apolar area buried upon folding ( $\Delta \Delta A_{pA_{U-F}}$ ) between each halophilic and non-halophilic

pair, have been calculated. According to the Wilcoxon and the *t*-test, none of the average values were significantly different from 0. Distribution of the differences in the SALTIN and OSMOL samples are reported in additional file 2.

#### Conserved hydrophobic contact area

All 15 structures collected from halophilic organisms have been structurally aligned with the non-halophilic counterparts, and the structurally conserved regions (SCR) were calculated. The conserved contacts between hydrophobic residues (CHC) were identified for each pair of structures, and were selected only those formed by evolutionary conserved residues. Then for each CHC, the apolar contact area formed by the pair of residues was calculated. The number of CHCs selected varies for each enzymatic family depending on the relative degree of conservation of residues in that family (Table 4).

**Table 4 Number of CHCs found in the two samples with the corresponding  $\Delta$ ACA.**

SALTIN HALOPHILE	Seq. lenght	NON HALOPHILE	Seq. lenght	Sequence identity(%)	No. CHC	OSMOL HALOPHILE	Seq. lenght	NON HALOPHILE	Seq. lenght	Sequence identity(%)	No. CHC
1DOI	128	1FXA	98	51	50	1NWZ	125	1MZU	129	46	55
1TJO	182	2VXX	192	36	36	1CNO	87	1ETP	190	44	26
2B5W	357	2CD9	366	30	154	1NML	326	3HQ6	345	47	157
2CC6	68	2V18	68	42	33	2VPN	316	3FXB	326	64	132
1ITK	731	2FXG	748	60	514	3IGN	177	3I5C	206	62	63
2AZ3	164	3B54	161	54	71	3BSM	413	2QJJ	402	40	187
2J5K	304	1Y6J	318	33	101	3IBM	167	3KGZ	156	66	65
3IFV	247	1RWZ	245	36	104						
Total <sup>a)</sup>					329.33						58.56
Average <sup>b)</sup>					-0.31						-0.08
t-test <sup>c)</sup>					<b>0.03</b>						0.52
Wilcoxon <sup>c)</sup>					<b>0.01</b>						0.63

Number of conserved hydrophobic contacts (CHCs) found in the two samples and the corresponding overall difference of apolar contact area ( $\Delta$ ACA)

a) Total  $\Delta$ ACA in the CHCs

b) Average  $\Delta$ ACA in the CHCs

c) Boldfaced digits indicate significant  $p$ -values

In SALTIN sample the average ACA in CHCs is significantly smaller in halophilic enzymes than non-halophilic counterparts (Table 4). At variance, OSMOL sample displays no significant difference. It has been also tested whether the differences of the CHC areas in the halophilic and non-halophilic counterparts correlated with sequence identity. No correlation was found (see additional file 3).

Statistics on the preferred amino acid exchanges at the SALTIN CHCs indicate that Ile  $\rightarrow$  Val is the most frequent replacement found in the direction mesophile  $\rightarrow$  halophile with a  $Z$ -score equal to 4.34. At variance, the most frequent exchange in the OSMOL CHCs is Leu  $\rightarrow$  Ile with  $Z$ -score 3.74.

## Discussion

Halophilic proteins have been extensively studied since a long time [10] and very much is known about their distinguishing characteristics, although several aspects not sufficiently elucidated of their molecular adaptation mechanisms still remain. Indeed, most of the published studies relied on comparative analyses of the characteristics of halophilic proteome and genomic sequences that provided general overview of the molecular signature of haloadaptation [15] but may not unravel subtle differences at the three-dimensional structural level. Likewise, three-dimensional structural comparisons were generally carried out within a single or a few protein families which may not be sufficiently representative of the halophilic proteome. Therefore, our work was aimed at contributing to the understanding of the shared molecular mechanisms of protein haloadaptation through a detailed and systematic comparison of a larger sample of several three-dimensional structures of halophilic and

non-halophilic proteins now available in the databanks. This approach highlights the adaptation strategies shared by the extreme halophilic proteins, besides the variations related to the peculiarities of each single family. Indeed, the extent to which individual proteins adapt to halophilic conditions varies, presumably due to their diverse characteristics and roles within the cell. For example, the malate dehydrogenase from *Salinibacter ruber* recently characterized [21] shares characteristics of a haloadapted archaeal enzyme and of non-haloadapted enzymes from other eubacterial species. Moreover, a proliferating cell nuclear antigen from *Haloferax volcanii* [16] displayed an unexpectedly low number of ion pairs at the monomer-monomer interface.

Proteins were taken both from microorganisms (mostly extreme halophilic *Archaea*) that utilize the "salt-in" strategy to cope with the strong environmental osmotic pressure, and from halophilic microorganisms using the intracellular accumulation of osmolytes and low cytoplasmatic salt concentration to counterbalance the external osmotic pressure. Proteins synthesized by the "salt-in" organisms are surrounded by a high-salt concentration environment and are in contact with it. At variance, intracellular proteins from the "osmolyte" cells are surrounded by a high concentration solution of compatible solutes which apparently do not interfere significantly with protein solubility, stability and activity [22]. Structural differences observed between the "salt-in" and the mesophilic homologous proteins (this set was referred to as SALTIN sample) were contrasted to those observed between the "osmolyte" and mesophilic pairs (here called OSMOL sample). The reasons for this comparative approach is twofold: the OSMOL sample

was the reference sample against which to test the structural differences observed in the SALTIN sample; at the same time we verified the presence of possible molecular adaptations in the OSMOL proteins due to their immersion into the osmolyte solution. As a further filter to eliminate possible noise due to non-normality distribution or paucity of data, two statistical tests (the parametric *t*-test and the non-parametric Wilcoxon test) were used to assess the significance of observed differences of structural parameters. Only differences defined significant according to both tests were deemed *bona fide* significant signals.

Starting from the surface features of the halophilic proteins, we demonstrated a significant reduction of apolar surface in the SALTIN sample and a parallel increase of the polar area formed by the side chain oxygen atoms, while the nitrogen side chain atoms lessen their contribution to the surface. These results reflect the well-known increase of surficial aspartate and glutamate residue frequency combined to the decrease of lysine residues [15,23]. It should be mentioned indeed that a reduction of hydrophobic exposed surface in the glucose dehydrogenase from *Haloferax mediterranei* has been demonstrated [24]. In this case however, reduction of apolar surface was caused by reduction of lysine residues. Thus, overall, the 2-fold reduction in the proportion of lysine residues in the sequence leads to a 4-fold reduction in the exposed hydrophobic accessible surface area contributed by the associated alkyl component of the lysine side chain.

We noticed that a significant decrease of apolar surface area in the protein synthesized by the halophilic organism can be observed in the OSMOL sample as well, paralleled by an ostensible tendency to increase the presence of side chain oxygen atoms (Table 2). Most of the compatible osmolytes produced by halophilic prokaryotes are neutral species, which contain charged groups (for example amino acid derivatives); therefore, it is not surprising that maintenance of solubility and stability of proteins at decimolar osmolyte concentration, required surface remodeling partly similar to that observed in the SALTIN sample, implying decrease of apolar surface. Indeed, several physical-chemical models have been proposed in the literature, according to which the charged sites of the stabilizing osmolytes interact with the oppositely charged polar areas of the macromolecule surface [25]. On the contrary, stabilizing osmolytes avoid contact with the polypeptide backbone. This property is called "osmophobic effect" [26] and it forces protein to fold *in vivo*, complementing the well known hydrophobic effect. However, the overall extent of surface modification is much smaller than that observed in the extreme halophilic proteins. Indeed, the examination of the surface atomic potential variations in the SALTIN

sample (Table 3) confirms that the electrostatic potential is significantly more negative in halophilic proteins rather in their mesophilic counterparts. Negative surface electrostatic potential has been demonstrated to be a characteristic factor of haloadaptation like, for example, in nucleoside diphosphate kinase from *Halobacterium salinarum* [27] or in an esterase from *Haloarcula marismortui* [19]. Thermal stability of the former protein indeed resulted to be strongly dependent on salt concentration, as predicted by theoretical studies [18]. The picture is different in the case of the OSMOL sample: the increase of the surface formed by side-chain oxygen atoms is not coupled to the decrease of nitrogen atoms: this account for the lack of an apparent modification of surficial potential. Recently, it has been demonstrated [17] by extensive site-directed mutagenesis on three protein domains (the halophilic 1A domain of the NAD-dependent DNA ligase N from *Haloferax volcanii*, the homologous domain from *E.coli*, and the mesophilic IgG binding domain of the protein L from *Streptococcus magnus*) that halophilicity of these proteins is directly related to a decrease in the solvent accessible surface. Authors stated that reduction of the solvent accessible area introduced upon mutation causes a progressive destabilization of the molecule, probably due to a reduction in the protein's hydrophobic effect. As a consequence, mutations increasing salt induced stabilization also destabilize the protein in the absence of salt, converting a mesophilic protein into an obligate halophilic form, a trend found in natural halophilic proteins.

However, we did not observe any significant reduction of the accessibility surface (data not shown) in the SALTIN and OSMOL halophilic proteins whereas we noticed the tendency, in the halophilic proteins of the SALTIN samples, to reduce the apolar contact area of the residues exposing more than 25% of their surface. Therefore one could argue that the reduction of hydrophobic effect be a key component of the haloadaptation mechanisms of surfaces.

However, surface property variations are not the only feature relevant to protein stability and solubility in high salt environments. Indeed, proper folding of the protein in these conditions requires the formation of appropriate hydrophobic interactions in the interior of the macromolecule. This is supposedly even more important during the early step of protein folding of the nascent polypeptide chain. In this perspective view, we analyzed the variations of apolar contact area in the hydrophobic core of the halophilic SALTIN and OSMOL surface compared to their mesophilic counterparts. No statistically significant difference could be found. Nonetheless, it should be noted that a consistent decrease of the apolar contact area in the halophilic proteins of the SALTIN sample can be observed for core residues at 0%

relative accessibility (Table 4). We believe this is indicative of a trend, absent in the OSMOL sample, related to haloadaptation of the SALTIN proteins. Similar trend can be observed at the subunit interfaces of extreme halophilic proteins.

Differences in apolar area buried upon folding ( $\Delta\Delta A_{pA_{U-F}}$ ) between each halophilic and non-halophilic pairs display overall no significant difference. This may suggest that, globally, the fraction of apolar area lost during folding is similar in the two cases (halophilic/non-halophilic) although the apolar surface available to the halophilic protein is smaller. This observation is reflected and supported by the significant reduction of exposed apolar area (Table 2) and apolar contact area (Table 4). It should be mentioned however that the difficult definition of unfolded state of a protein [28] should recommend caution in the interpretation of these results. Moreover, in this specific problem, electrostatic repulsion of all the negative charges can render the calculation of the solvent accessible surface area of unfolded state even more inaccurate.

It is well known that high NaCl concentration strengthens hydrophobic interactions [29] and recently the thermodynamics of this effect has been analyzed in synthetic polymers [30]. Hydrophobicity is the main driving force of protein folding [31]; it must be finely balanced to confer proteins proper flexibility and stability compatible with their function. To maintain the magnitude of hydrophobic interactions within useful intervals in a high salt environment, extent of protein interior hydrophobic surface should be reduced in order to lessen the change in solvated hydrophobic areas. Interestingly, it was demonstrated [30] that a potent osmolyte, trimethylamine oxide, have a negligible effect on the strength of hydrophobic interactions. This finding suggests that hydrophobic interactions in proteins of "osmolyte" microorganisms need not to be significantly altered.

This picture is even clearer when inspecting and contrasting the conserved hydrophobic contacts in the proteins belonging to the SALTIN and OSMOL sample. The CHCs [18] are sites in the interior of the protein in correspondence of which there is a hydrophobic contact conserved during divergent evolution. Evolutionary conservation suggests that the contact be essential for folding and/or stability of the protein. The average apolar contact area at the CHCs is significantly smaller in halophilic proteins than in their mesophilic counterparts. As expected, OSMOL sample display no difference (Table 4). Accordingly, the most frequent residue exchange in the SALTIN CHCs is the replacement of Ile with Val in halophilic proteins which reduces the hydrophobic volume buried in the core. Indeed, the effect of such substitution is the loss of a methyl group in the

hydrophobic nucleus of the protein. This is expected to reduce the hydrophobic interactions, although conformational strain can also compound the influence on the stability of the protein [32]. As expected, the most frequent residue substitution of the OSMOL CHCs, namely Leu→Ile, does not imply any reduction in the number of side-chain carbon atoms. As an example, Figure 3 reports two cases taken from the dataset utilized in the work in which a CHC Ile was replaced by Val in the halophilic protein.

## Conclusions

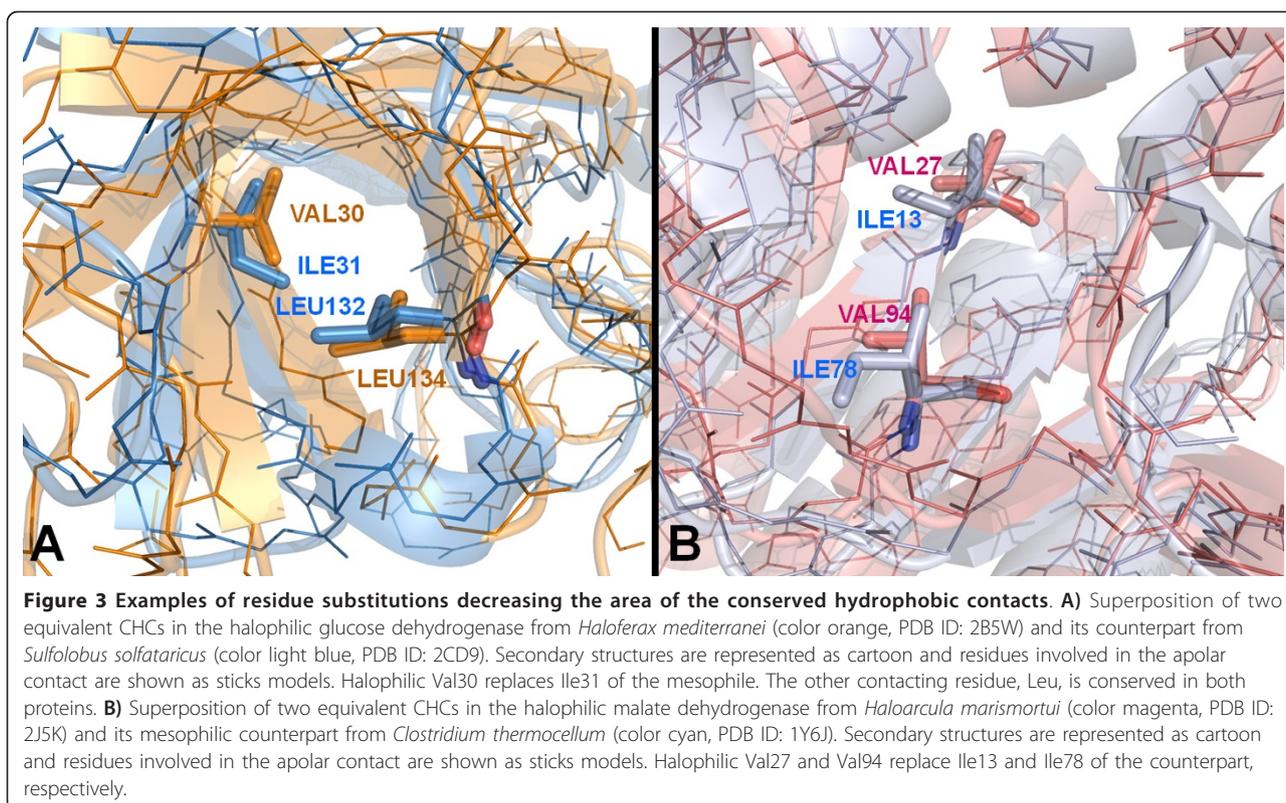
The present analysis suggests that a shared haloadaptation strategy of proteins in the presence of molar salt concentration, but not in the presence of osmolytes, necessitates a weakening of the hydrophobic interactions in general, and in particular at the level of core and conserved hydrophobic contacts. Weakening of these interactions counterbalances their strengthening by the presence of salts in solution and may help the structure preventing aggregation and/or loss of function in hypersaline environments. Indeed, decreasing hydrophobicity makes halophilic proteins unstable in low-concentration salt solutions and may in part explain the request of the halophilic proteins for high salt concentrations. To complete the picture, the destabilization of halophilic proteins at low-salt concentration due to the strong electrostatic repulsion should be considered [18,33]. Shrinking of hydrophobic contacts must be even more critical for the early stages of folding when intramolecular hydrophobic nuclei must correctly form to guide the polypeptide through the folding funnel to the native state.

Considering also the significant increase of biotechnology applications of halophiles, the comprehension of the multifaceted etiology of halophilicity (including the electrostatic factors) can provide the theoretical basis for the engineering of proteins of great interest because stable at concentrations of salts that cause the denaturation or aggregation of the majority of macromolecules.

## Methods

### Selection of protein structures

The crystallographic structures of the available high salt concentration active enzymes were found in PDB [34]. The search was carried out using the keywords: "halophil", "haloarc", "halobacter" and the like. For each of the obtained structures, it was verified that the source organism was present in a previously compiled list of halophilic organisms, based on the website XBASE [35] <http://www.xbase.ac.uk/> and in the literature. The selected organisms have optimal growth conditions that support the presence of genuine adaptation to high salt concentration environments. In fact, those organisms



requiring for optimum growth concentration of NaCl < 0.5 M and/or temperatures < 20°C or > 55°C, were excluded from the data sample. Moreover, the organisms were divided into two groups corresponding to different survival strategies: in the first group are the archaeal halophiles using the “salt-in” strategy (which will be referred to as SALTIN sample); in the latter the *Eubacteria* using the “osmolytes” strategy (OSMOL sample) [36].

To select only halophilic structures matching specific criteria and to eliminate redundancy, we used the PISCES server <http://dunbrack.fccc.edu/PISCES.php>[37], which enables the creation of a list of structures that meet arbitrary thresholds, starting from an initial user-defined larger sample. The parameters used were: maximum percentage identity 90%, maximum resolution 2.5 Å, maximum R-value 0.3, minimum chain length 50 residues, and utmost chain length 1000 residues. Homologous structures from non-halophilic organisms were retrieved from PDB by means of the program PSI-BLAST [38]. To ensure structural homology, only sequences sharing  $\geq 30\%$  and  $\leq 90\%$  residue identity to the halophilic counterpart were considered. In cases where more counterparts were available for each halophilic protein, only the homolog with the highest percentage identity to the protein of interest was considered to minimize possible phylogenetic drift

effects on the statistics. At the end of the procedure, 15 pairs of halophilic and non-halophilic proteins were collected, 8 of which corresponded to organisms that adopt the “salt-in” strategy and constitute the SALTIN sample, and 7 that adopt the “osmolytes” strategy and compose the OSMOL sample. In cases where the protein structures are used in their biological unit form, the pairs are reduced to 11 because in 4 cases there is no correspondence between the quaternary structure of homologous halophilic and non-halophilic proteins. Although Table 1 shows that the two proteins of the pair 1CNO-1ETP have a different quaternary structure, this pair was included in the analysis. In fact, since 1ETP is a fusion protein of two polypeptide chains each equivalent to the monomer of 1CNO, the two quaternary forms are overall similar.

#### Apolar contact area (ACA)

The apolar contact area is the area of contact between coupled non-polar atoms in a protein. It was calculated for different structural environments: nucleus, interface between monomers and protein surface. The program utilized was Pdb\_np\_cont [39], which calculates the area of contact between coupled non-polar atoms starting from a standard PDB file. Briefly, this method is based on the classification of points located on a sphere of interaction radius, surrounding each non-polar atom.

The interaction radius is the van der Waals radius of each atom type, plus the radius of a water molecule. The output of this program was utilized to calculate the pairwise residue contact areas for every possible pair of residues belonging to the structures analysed.

Individual protein chains were used for analysis of the protein core, while analyses of the protein subunit interfaces were carried out with structures in their quaternary assembly. Only the 9 pairs of proteins possessing an oligomeric biological unit were used in this case. Protein interfaces consists of those residues making apolar contacts to another protein chain, as defined by Pdb\_np\_cont [38].

A residue was assigned to the structural environment “nucleus” if its relative solvent accessibility, calculated with the program NACCESS [40], is less than four different alternative thresholds (0%, 3%, 5% and 9%). A residue is instead assigned to the protein surface if its relative solvent accessibility is greater than an arbitrarily fixed threshold. To this end, four different alternative thresholds (25%, 40%, 55% and 70%) were used.

For each structural environment considered the per-residue mean apolar contact area was calculated by dividing the total apolar contact area by the number of residues belonging to that structural environment.

Conserved hydrophobic contacts (CHCs) were identified in each pair of homologous enzymes through the combined use of the web tools CE-MC [41] for calculating the structural alignment, SCR\_FIND for identification of structurally conserved regions and CHC\_FIND for the detection of the CHCs and their apolar contact area [42]. SCRs were defined as regions displaying a similar local conformation, with a mean positional RMSD of the equivalent  $\alpha$ -carbon positions of the structures superposed  $\leq 3.0$  Å, lacking indels and composed of at least three consecutive residues. A CHC is defined as a conserved hydrophobic contact formed between two residues belonging to an SCR, in members belonging to a family or superfamily of proteins [42]. We considered only those contacts that are formed between residues distant at least three positions along the sequence and residues that are evolutionarily conserved according to the server CONSURF [43], *i.e.* belonging to the classes 7, 8 and 9. These three classes contain the most conserved amino acid positions from a total of nine equally sized categories of relative degree of conservation.

#### Preferred amino acid substitution in CHCs

Amino acid substitutions of residues involved in the formation of structurally and evolutionarily conserved hydrophobic contacts between halophilic and non-halophilic proteins were determined by analysing the alignment of the SCRs of each pair. For each residue  $X$ ,

belonging to a non-halophilic protein and involved in making CHCs,  $aa_{X \rightarrow Y}$  was defined as the number of times  $X$  is substituted by the residue  $Y$  of the halophilic protein. Likewise,  $aa_{Y \rightarrow X}$  is defined. Therefore, a substitution matrix can be obtained by computing the difference between  $aa_{X \rightarrow Y}$  and  $aa_{Y \rightarrow X}$  over the whole dataset of protein pairs  $k$ , according to:

$$C^S = \sum_k \left( \sum aa_{X \rightarrow Y} - \sum aa_{Y \rightarrow X} \right) \quad (1)$$

where  $C^S$  is the element of the substitution matrix.

The mean and standard deviation of the overall exchange matrix were determined; the significance  $R_{XY}$  of the exchange  $X \rightarrow Y$  was then calculated by dividing the difference between  $C^S$ , and the overall matrix mean  $\bar{C}$  by the standard deviation  $\sigma$ :

$$R_{XY} = \frac{C^S - \bar{C}}{\sigma} \quad (2)$$

$R_{XY}$  values  $\geq 3.0$  standard deviations (corresponding to a probability  $P \leq 0.01$  that the observed difference was obtained by chance) from the mean value were considered statistically significant [18].

#### Accessible surface area (ASA) analysis

The ASA is the area of the protein surface which can be in contact with the solvent. The total apolar (and the complementary polar) component of the entire accessible surface area of each protein in its quaternary form was calculated with the server GETAREA <http://curie.utmb.edu/getarea.html> with default settings [44]. The contribution of different atomic types to the polar area, namely oxygens, side-chain oxygens, nitrogens and side-chain nitrogens was also investigated. To compare the different surface areas of halophilic and non-halophilic proteins, they were normalized by division by the total accessible surface area of the corresponding protein.

The differences between the fraction of apolar accessible surface area in the unfolded and folded form of each protein considered ( $\Delta ApA_{U-F}$ ) were calculated through the web server <http://roselab.jhu.edu/utis/unfolded.html> [28]. The differences between the  $\Delta ApA_{U-F}$  in each halophilic protein and in its corresponding homolog ( $\Delta \Delta ApA_{U-F}$ ) were calculated and statistically tested.

#### Electrostatic potential calculation

The surface electrostatic potential of the proteins in their quaternary form was calculated using the program DELPHI [45]. Salt concentration was set equal to 0, since identical environmental conditions can better delineate differences between the electrostatic potential of the halophilic and non-halophilic homolog. Internal

and solvent dielectric constants were set to 4.0 [46] and 80.0 respectively. The other parameters used were set to the default values: grid scale = 1.2, box fill = 60%, probe radius = 1.4 Å, and van der Waals surface. To compare the potential of halophilic and non-halophilic proteins of different lengths, the average atomic potential (AAP) was calculated dividing the total electrostatic potential by the total number of atoms.

### Statistical tests

Whenever possible, differences between the structural properties considered were calculated within the two samples: the SALTIN halophilic and non-halophilic homologs, and the OSMOL halophilic and non-halophilic pairs. Differences of structural properties are denoted by  $\Delta$ : for example,  $\Delta$ ACA indicates halophilic minus non-halophilic apolar contact areas.

The  $\Delta$ s between the structural properties of halophilic enzymes and their corresponding non-halophilic counterparts were tested within samples using two statistical tests, a paired *t*-test and a non-parametric Wilcoxon signed-rank test [47]. In the former case the null hypothesis is that the average  $\Delta$  was 0 at 0.05 *p*-value while in the latter case the null hypothesis is that the median of the  $\Delta$ s was 0 at 0.05 *p*-value. The parametric *t*-test assumes that the tested data come from a normal distribution, while the Wilcoxon test it is less restrictive since it does not require such a condition. To enhance robustness of the conclusions drawn from the structural comparisons, only differences resulting significant from both tests were deemed *bona fide* significant results.

### Additional material

**Additional file 1: Table S1 -  $\Delta$ ASA in the SALTIN and OSMOL samples at conserved residues.** Differences of fractional accessibility surface area ( $\Delta$ ASA) in the SALTIN and OSMOL samples for different class of atoms. The differences are between the surface areas calculated in the halophilic protein and the corresponding areas in the non-halophilic counterpart. The calculations were performed considering the proteins in their quaternary structure. Only residues identically conserved in the two proteins were considered in the calculations.

**Additional file 2: Additional Figure 1 -  $\Delta$ ApA<sub>U-F</sub>.** Histograms reporting the  $\Delta$ ApA<sub>U-F</sub> in the SALTIN and OSMOL samples. The  $\Delta$ ApA<sub>U-F</sub> values were calculated as the difference between the fraction of exposed apolar area lost during folding of the halophilic protein and the fraction lost by the corresponding non-halophilic homolog. Further details are reported in the "Methods" section of the main text.

**Additional file 3: Additional Figure 2 - Correlation between single  $\Delta$ ACA and pairwise percentage identity.** Graph reporting the difference between the area of each halophilic CHC and that of the corresponding non-halophilic CHC ( $\Delta$ ACA) versus the pairwise sequence percentage identity (%id) for the SALTIN and OSMOL samples.

### List of abbreviations

$\Delta$ : prefix indicating the difference between the property measured in the halophilic and in the non-halophilic homologous proteins; AAP: average

atomic potential; ACA: apolar contact area; ApA: apolar accessible area;  $\Delta$ ApA<sub>U-F</sub>: difference between ApA in the unfolded and folded state; ASA: solvent accessible surface area; CHC: conserved hydrophobic contacts; OSMOL: halophilic organisms adopting the "osmolytes" strategy; RMSD: Root mean square deviation; SALTIN: halophilic organisms adopting the "salt-in" strategy; SCR: structurally conserved region.

### Acknowledgements

This work has been partially supported by the funds from the Italian "Ministero dell'Istruzione, dell'Università e della Ricerca". This work will be submitted by AS in partial fulfillment of the requirements of the degree of "Dottorato di Ricerca in Biochimica" at Sapienza, Università di Roma. This paper is dedicated to Prof. Francesco Bossa on the occasion of his 70<sup>th</sup> birthday.

### Authors' contributions

AS and MP collected the data, carried out the calculations and analysed results. AP conceived the study, helped to analyse results and to correct the manuscript. SP analysed the results, drafted the manuscript and participated in study design and coordination. All authors read and approved the final manuscript and declare no conflict of interest.

Received: 10 August 2011 Accepted: 22 December 2011  
Published: 22 December 2011

### References

1. Egorova K, Antranikian G: Industrial relevance of thermophilic *Archaea*. *Curr Opin Microbiol* 2005, **8**:649-655.
2. Pikuta EV, Hoover RB, Tang J: Microbial extremophiles at the limits of life. *Crit Rev Microbiol* 2007, **33**:183-209.
3. Oren A: Saltern evaporation ponds as model systems for the study of primary production processes under hypersaline conditions. *Aquat Microb Ecol* 2009, **56**:193-204.
4. Ebrahimie E, Ebrahimi M, Sarvestani NR, Ebrahimi M: Protein attributes contribute to halo-stability, bioinformatics approach. *Saline Systems* 2011, **7**.
5. Feller G: Protein stability and enzyme activity at extreme biological temperatures. *J Phys Condens Matter* 2010, **22**.
6. Oren A: Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline Systems* 2008, **4**.
7. Gunde-Cimerman N, Ramos J, Plemenitas A: Halotolerant and halophilic fungi. *Mycol Res* 2009, **113**:1231-1241.
8. Roberts MF: Organic compatible solutes of halotolerant and halophilic microorganisms. *Saline Systems* 2005, **1**.
9. Zaccai G, Eisenberg H: Halophilic proteins and the influence of solvent on protein stabilization. *Trends Biochem Sci* 1990, **15**:333-337.
10. Madern D, Ebel C, Zaccai G: Halophilic adaptation of enzymes. *Extremophiles* 2000, **4**:91-98.
11. Mevarech M, Frolow F, Gloss LM: Halophilic enzymes: protein with a grain of salt. *Biophys Chem* 2000, **86**:155-164.
12. Wright DB, Banks DD, Lohman JR, Hilsenbeck JL, Gloss LM: The effect of salts on the activity and stability of *Escherichia coli* and *Haloflex volcanii* dihydrofolate reductases. *J Mol Biol* 2002, **18**:327-344.
13. Bergqvist S, Williams MA, O'Brien R, Ladbury JE: Halophilic adaptation of protein-DNA interactions. *Biochem Soc Trans* 2003, **31**:677-680.
14. Fukuchi S, Yoshimune K, Wakayama M, Moriguchi M, Nishikawa K: Unique amino acid composition of proteins in halophilic bacteria. *J Mol Biol* 2003, **327**:347-357.
15. Paul S, Bag SK, Das S, Harvill ET, Dutta C: Molecular signature of hypersaline adaptation: insights from genome and proteome composition of halophilic prokaryotes. *Genome Biol* 2008, **9**.
16. Winter JA, Christofi P, Morroll S, Bunting KA: The crystal structure of *Haloflex volcanii* proliferating cell nuclear antigen reveals unique surface charge characteristics due to halophilic adaptation. *BMC Struct Biol* 2009, **9**.
17. Tadeo X, López-Méndez B, Trigueros T, Laín A, Castaño D, Millet O: Structural Basis for the Aminoacid Composition of Proteins from Halophilic *Archaea*. *PLoS Biology* 2009, **7**.
18. Elcock AH, McCammon JA: Electrostatic contributions to the stability of halophilic proteins. *J Mol Biol* 1998, **280**:731-748.

19. Müller-Santos M, de Souza EM, Pedrosa Fde O, Mitchell DA, Longhi S, Carrière F, Cnaan S, Krieger N: **First evidence for the salt-dependent folding and activity of an esterase from the halophilic *Archaea Haloarcula marismortui***. *Biochim Biophys Acta* 2009, **1791**:719-729.
20. Paiardini A, Sali R, Bossa F, Pascarella S: **"Hot cores" in proteins: comparative analysis of the apolar contact area in structures from hyper/thermophilic and mesophilic organisms**. *BMC Struct Biol* 2008, **8**.
21. Coquelle N, Talon R, Juers DH, Girard E, Kahn R, Madern D: **Gradual adaptive changes of a protein facing high salt concentrations**. *J Mol Biol* 2010, **404**:493-505.
22. Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN: **Living with water stress: evolution of osmolyte systems**. *Science* 1982, **217**:1214-1222.
23. Rhodes ME, Fitz-Gibbon ST, Oren A, House CH: **Amino acid signatures of salinity on an environmental scale with a focus on the Dead Sea**. *Environ Microbiol* 2010, **12**:2613-2623.
24. Britton KL, Baker PJ, Fisher M, Ruzheinikov S, Gilmour DJ, Bonete MJ, Ferrer J, Pire C, Esclapez J, Rice DW: **Analysis of protein solvent interactions in glucose dehydrogenase from the extreme halophile *Haloferax mediterranei***. *Proc Natl Acad Sci USA* 2006, **103**:4846-4851.
25. Street TO, Bolen DW, Rose GD: **A molecular mechanism for osmolyte-induced protein stability**. *Proc Natl Acad Sci USA* 2006, **103**:13997-14002.
26. Bolen DW, Baskakov IV: **The osmophobic effect: natural selection of a thermodynamic force in protein folding**. *J Mol Biol* 2001, **310**:955-963.
27. Besir H, Zeth K, Bracher A, Heider U, Ishibashi M, Tokunaga M, Oesterhelt D: **Structure of a halophilic nucleoside diphosphate kinase from *Halobacterium salinarum***. *FEBS Lett* 2005, **579**:6595-6600.
28. Creamer TP, Srinivasan R, Rose GD: **Modeling unfolded states of peptides and proteins**. *Biochemistry* 1995, **34**:16245-16250.
29. Ghosh T, Kalra A, Garde S: **On the salt-induced stabilization of pair and many-body hydrophobic interactions**. *J Phys Chem B* 2005, **109**:642-651.
30. Athawale MV, Sarupria S, Garde S: **Enthalpy-entropy contributions to salt and osmolyte effects on molecular-scale hydrophobic hydration and interactions**. *J Phys Chem B* 2008, **112**:5661-5670.
31. Chandler D: **Interfaces and the driving force of hydrophobic assembly**. *Nature* 2005, **437**:640-647.
32. Ventura S, Vega MC, Lacroix E, Angrand I, Spagnolo L, Serrano L: **Conformational strain in the hydrophobic core and its implications for protein folding and design**. *Nat Struct Biol* 2002, **9**:485-493.
33. Arakawa T, Tokunaga M: **Electrostatic and hydrophobic interactions play a major role in the stability and refolding of halophilic proteins**. *Protein Pept Lett* 2004, **11**:125-132.
34. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE: **The Protein Data Bank**. *Nucleic Acids Res* 2000, **28**:235-242.
35. Chaudhuri RR, Pallen MJ: **xBASE, a collection of online databases for bacterial comparative genomics**. *Nucleic Acids Res* 2006, **34**:335-337.
36. Dennis PP, Shimmin LC: **Evolutionary divergence and salinity-mediated selection in halophilic *Archaea***. *Microbiol Mol Biol Rev* 1997, **61**:90-104.
37. Wang G, Dunbrack RL Jr: **PISCES: a protein sequence culling server**. *Bioinformatics* 2003, **19**:1589-1591.
38. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs**. *Nucleic Acids Res* 1997, **25**:3389-3402.
39. Drablos F: **Clustering of non-polar contacts in proteins**. *Bioinformatics* 1999, **15**:501-509.
40. Hubbard SJ, Thornton JM: **Naccess, Computer Program**. Department of Biochemistry and Molecular Biology, University College, London; 1993.
41. Guda C, Lu S, Scheeff ED, Bourne PE, Shindyalov IN: **CE-MC: a multiple protein structure alignment server**. *Nucleic Acids Res* 2004, **32**:100-103.
42. Paiardini A, Bossa F, Pascarella S: **CAMPO, SCR\_FIND and CHC\_FIND: a suite of web tools for computational structural biology**. *Nucleic Acids Res* 2005, **33**:50-55.
43. Glaser F, Pupko T, Paz I, Bell RE, Bechor-Shental D, Martz E, Ben-Tal N: **ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information**. *Bioinformatics* 2003, **19**:163-164.
44. Frackiewicz R, Braun W: **Exact and Efficient Analytical Calculation of the Accessible Surface Areas and Their Gradients for Macromolecules**. *J Comp Chem* 1998, **19**:319-333.
45. Rocchia W, Alexov E, Honig B: **Extending the Applicability of the Nonlinear Poisson-Boltzmann Equation: Multiple Dielectric Constants and Multivalent Ions**. *J Phys Chem B* 2001, **105**:6507-6514.
46. Gilson MK, Honig BH: **The dielectric constant of a folded protein**. *Biopolymers* 1986, **25**:2097-2119.
47. Wilcoxon F: **Individual comparisons by ranking methods**. *Biometrics Bulletin* 1945, **1**:80-83.

doi:10.1186/1472-6807-11-50

**Cite this article as:** Siglioccolo et al.: Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface. *BMC Structural Biology* 2011 **11**:50.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

